

Data Management for Big Data (DMBD)
and
Advanced Database Systems for Big Data (ADBS4BD)

Dario Della Monica

April 3, 2023

Teacher's information

Dario Della Monica

Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF)
University of Udine

- ▶ **email:** dario.dellamonica@uniud.it
- ▶ **web:** <http://users.dimi.uniud.it/~dario.dellamonica/>
 - ▶ [Home] >> Teaching >> Advanced Database Systems For Big Data
 - ▶ [Home] >> Teaching >> Data Management for Big Data
- ▶ **office hours:** Tuesday, 16:30-18:30 (or by appointment)
 - ▶ appointment is always preferable
- ▶ more info (**cell**, **skype**, ...) on my webpage

Course structure

These lectures are part of 2 courses

- ▶ Advanced Database Systems for Big Data (ADBS4BD) – 48h, 6 CFU
 - ▶ Master Degree in Artificial Intelligence & Cybersecurity, University of Udine
 - ▶ Master Degree in Computer Science, University of Udine
- ▶ Data Management for Big Data (DMBD) – 72h, 9 CFU
 - ▶ Master Degree in Data Science and Scientific Computing
University of Udine and University of Trieste, Italy

Overview of the lectures

- ▶ Part 0: Fundamentals of database systems (24h, 3 CFU)
 - ▶ already delivered by Adriano Peron
 - ▶ only for DMBD course
- ▶ Part 1: Advanced database models, languages, and systems (24h, 3 CFU)
 - ▶ delivered by Dario Della Monica
 - ▶ for both DMBD and ADBS4BD courses
- ▶ Part 2: Data analysis and big data (24h, 3 CFU)
 - ▶ to be delivered by Adriano Peron
 - ▶ for both DMBD and ADBS4BD courses

Overview of Part 1

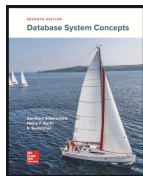
Advanced database models, languages, and systems

- ▶ Query processing and optimization
 - ▶ Query processing
 - ▶ algorithms for evaluating operations (e.g., selection, join)
 - ▶ Cost-based optimization and heuristics
 - ▶ choice of “best” evaluation plan
- ▶ Distributed database architectures
 - ▶ An introduction to distributed DBMS
 - ▶ Distributed database design (fragmentation and allocation)
 - ▶ Distributed query processing and optimization
 - ▶ Transaction management in distributed databases
 - ▶ distributed concurrency control (two-phase locking – 2PL)
 - ▶ distributed DBMS reliability (two-phase commit – 2PC)

Material about Part 1

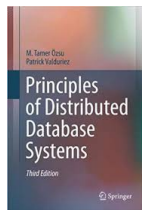
- Database System Concepts
Silberschatz, Korth, Sudarshan
7th Edition (or 6th Edition)
McGraw-Hill, 2020

<https://www.db-book.com/db7/index.html>
(extra contents: original slides, solutions, etc)



- Principles of Distributed Database Systems
Özsu and Valduriez
3rd Edition
Springer, 2011

ISBN: 978-1-4939-4174-2/978-1-4419-8834-8 (e-book)
<https://extras.springer.com/2011/978-1-4419-8833-1.zip>
(original slides)



- other material (e.g., slides) will be made available on my webpage